GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 09:23:04 ; Search time 592 Seconds
(without alignments)
10369.125 Million cell updates/sec

Title: US-09-494-297-1
Perfect score: 2274
Sequence: 1 atgaaaaaaacaaggtttcc.....ggataagaaaacatgactag 2274

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2271	99.9	2286	24	ABN69678	Streptococcus poly
c	2	75	3.3	159	24	ABN69679	Streptococcus poly
	3	56.6	2.5	4677	21	AAA70259	Plasmodium falcipa
	4	56.6	2.5	32392	24	ABL56203	AmEPV genome fragm
c	5	56.2	2.5	34688	24	ABQ67060	Human angiogenesis
	6	56	2.5	11922	21	AAA70187	Plasmodium falcipa
	7	53.8	2.4	5940	21	AAA70105	Plasmodium falcipa
	8	53	2.3	494	23	ABV10021	Human prostate exp
	9	53	2.3	335913	22	AAI61371	Soybean 240017 reg
	10	53	2.3	335913	22	AAI61372	Soybean 240017 reg
c	11	52.8	2.3	8210	24	ABL70331	Chemically treated
c	12	52.8	2.3	8210	24	AAS61282	Human gene regulat
c	13	52.8	2.3	8210	24	ABK31380	Signal transductio
c	14	52.4	2.3	6106	22	AAS46429	Tumour suppressor
c	15	52.4	2.3	6106	24	ABK40031	Human chemically p
c	16	52.4	2.3	6106	24	ABL33472	Human immune syste
c	17	52.2	2.3	50000	24	ABL56202	AmEPV genome fragm
c	18	52	2.3	7195	22	AAS45325	Chemically pretrea
c	19	52	2.3	7195	24	ABK28166	DNA transcription
c	20	52	2.3	37515	24	ABQ66998	Human angiogenesis
	21	51.4	2.3	1998	21	AAA70212	Plasmodium falcipa
	22	51.2	2.3	1527	21	AAA70121	Plasmodium falcipa
	23	51.2	2.3	3399	17	AAT05868	Chicken leucocytoz
c	24	51.2	2.3	4985	24	ABQ75107	Anopheles gambiae
c	25	51	2.2	8056	25	ABZ10100	Haematopoietic cel
	26	50.6	2.2	4590	22	AAH24065	Yeast AOD9604-asso
c	27	50.6	2.2	6815	22	AAS45345	Chemically pretrea
c	28	50.6	2.2	6815	24	ABL32671	Human immune syste
c	29	50.6	2.2	6815	24	ABK28176	DNA transcription
c	30	50.6	2.2	7819	24	ABL33953	Human immune syste
c	31	50.6	2.2	7819	24	ABL34607	Human metastasis a
c	32	50.2	2.2	15016	20	AAX99560	Nucleic acid seque
c	33	50.2	2.2	19087	24	ABL32793	Human immune syste
	34	50	2.2	2000	24	ABZ14930	Arabidopsis thalia
c	35	50	2.2	6381	24	ABL70243	Chemically treated
c	36	50	2.2	6381	24	ABL32966	Human immune syste
c	37	50	2.2	6381	24	ABL34518	Human metastasis a
	38	50	2.2	640681	24	ABA92787	Buchnera sp. genom
c	39	49.8	2.2	50000	24	ABL55643	AmEPV genome fragm
c	40	49.8	2.2	50000	24	ABL56201	AmEPV genome fragm
	41	49.6	2.2	3095	11	AAQ03875	Sequence encoding
c	42	49.6	2.2	12705	24	ABL32148	Human immune syste
	43	49.4	2.2	15251	24	ABQ76622	C. albicans BAX-as
c	44	49.2	2.2	34548	24	ABL70603	Chemically treated
c	45	49	2.2	16724	24	ABL70259	Chemically treated

ALIGNMENTS

RESULT 1
ABN69678
ID ABN69678 standard; DNA; 2286 BP.
XX
AC ABN69678;
XX
DT 01-JUL-2002 (first entry)
XX
DE Streptococcus polynucleotide SEQ ID NO 7269.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX
OS Streptococcus pyogenes.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB04789.
XX
PR 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
PI Tettelin H;

XX
DR WPI; 2002-352536/38.
DR P-PSDB; ABP29047.
XX
PT New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
PS Claim 7; Page 3879; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyrogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
SQ Sequence 2286 BP; 831 A; 347 C; 447 G; 661 T; 0 other;

Query Match 99.9%; Score 2271; DB 24; Length 2286;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGAAAAAAACAAGGTTTCCAAATAAGCTTAATACTCTTAATACTCAAAGGGTATTAAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 ATGAAAAAAACAAGGTTTCCAAATAAGCTTAATACTCTTAATACTCAAAGGGTATTAAGT 75

Qy         61 AAAAACTCAAAACGATTTACTGTCACTTTAGTGGGAGTCTTTTTAATGATCTTCGCTTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 AAAAACTCAAAACGATTTACTGTCACTTTAGTGGGAGTCTTTTTAATGATCTTCGCTTTG 135

Qy        121 GTAACTTCCATGGTTGGTGCTAAGACTGTTTTTGGTTTAGTAGAATCCTCGACGCCAAAC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 GTAACTTCCATGGTTGGTGCTAAGACTGTTTTTGGTTTAGTAGAATCCTCGACGCCAAAC 195

Qy        181 GCAATAAATCCAGATTCAAGTTCGGAATACAGATGGTATGGATATGAATCTTATGTAAGA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 GCAATAAATCCAGATTCAAGTTCGGAATACAGATGGTATGGATATGAATCTTATGTAAGA 255

Qy        241 GGGCATCCATATTATAAACAGTTTAGAGTAGCACACGATTTAAGGGTTAACTTAGAAGGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 GGGCATCCATATTATAAACAGTTTAGAGTAGCACACGATTTAAGGGTTAACTTAGAAGGA 315

Qy        301 AGTAGAAGTTATCAAGTTTATTGCTTTAATTTAAAGAAAGCATTTCCTCTCGGATCAGAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 AGTAGAAGTTATCAAGTTTATTGCTTTAATTTAAAGAAAGCATTTCCTCTCGGATCAGAT 375

Qy        361 AGTAGTGTTAAAAAGTGGTATAAAAAACATGATGGAATCTCTACAAAATTTGAAGATTAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 AGTAGTGTTAAAAAGTGGTATAAAAAACATGATGGAATCTCTACAAAATTTGAAGATTAT 435

Qy        421 GCGATGAGCCCTAGAATTACGGGAGATGAGCTAAATCAGAAGTTACGAGCTGTTATGTAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 GCGATGAGCCCTAGAATTACGGGAGATGAGCTAAATCAGAAGTTACGAGCTGTTATGTAT 495

Qy        481 AATGGACATCCACAAAATGCCAATGGTATTATGGAAGGCTTGGAACCCTTGAATGCTATC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 AATGGACATCCACAAAATGCCAATGGTATTATGGAAGGCTTGGAACCCTTGAATGCTATC 555

Qy        541 AGAGTTACACAAGAGGCGGTATGGTACTATTCTGATAATGCTCCTATTTCTAATCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 AGAGTTACACAAGAGGCGGTATGGTACTATTCTGATAATGCTCCTATTTCTAATCCAGAT 615

Qy        601 GAAAGTTTTAAAAGGGAGTCAGAAAGTAACTTGGTTAGTACTTCTCAATTATCTTTGATG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 GAAAGTTTTAAAAGGGAGTCAGAAAGTAACTTGGTTAGTACTTCTCAATTATCTTTGATG 675

Qy        661 CGTCAAGCTTTGAAGCAACTGATTGATCCGAATTTGGCAACTAAAATGCCAAAACAAGTT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 CGTCAAGCTTTGAAGCAACTGATTGATCCGAATTTGGCAACTAAAATGCCAAAACAAGTT 735

Qy        721 CCGGATGATTTTCAGCTAAGTATTTTTGAGTCTGAGGACAAGGGAGATAAATATAATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 CCGGATGATTTTCAGCTAAGTATTTTTGAGTCTGAGGACAAGGGAGATAAATATAATAAA 795

Qy        781 GGATACCAAAATCTTTTGAGTGGTGGTTTAGTTCCTACTAAACCACCAACTCCAGGAGAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GGATACCAAAATCTTTTGAGTGGTGGTTTAGTTCCTACTAAACCACCAACTCCAGGAGAC 855

Qy        841 CCACCAATGCCTCCAAATCAACCTCAAACGACTTCAGTACTTATTAGAAAGTATGCTATA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 CCACCAATGCCTCCAAATCAACCTCAAACGACTTCAGTACTTATTAGAAAGTATGCTATA 915

Qy        901 GGTGATTACTCTAAATTGCTTGAAGGTGCAACATTACAGTTGACAGGGGATAACGTGAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 GGTGATTACTCTAAATTGCTTGAAGGTGCAACATTACAGTTGACAGGGGATAACGTGAAT 975

Qy        961 AGTTTTCAAGCGAGAGTGTTTAGCAGTAATGATATTGGAGAAAGAATTGAACTATCAGAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 AGTTTTCAAGCGAGAGTGTTTAGCAGTAATGATATTGGAGAAAGAATTGAACTATCAGAT 1035

Qy       1021 GGAACTTATACTTTAACTGAATTGAATTCTCCAGCTGGTTATAGTATCGCAGAGCCAATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 GGAACTTATACTTTAACTGAATTGAATTCTCCAGCTGGTTATAGTATCGCAGAGCCAATC 1095

Qy       1081 ACTTTTAAGGTTGAAGCTGGCAAAGTGTATACTATTATTGATGGAAAACAGATTGAAAAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 ACTTTTAAGGTTGAAGCTGGCAAAGTGTATACTATTATTGATGGAAAACAGATTGAAAAT 1155

Qy       1141 CCCAATAAAGAGATAGTAGAGCCTTACTCAGTAGAAGCATATAATGATTTTGAAGAATTT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1156 CCCAATAAAGAGATAGTAGAGCCTTACTCAGTAGAAGCATATAATGATTTTGAAGAATTT 1215

Qy       1201 AGCGTTTTAACTACACAAAACTATGCAAAATTTTATTATGCAAAAAATAAAAATGGAAGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1216 AGCGTTTTAACTACACAAAACTATGCAAAATTTTATTATGCAAAAAATAAAAATGGAAGT 1275

Qy       1261 TCACAGGTTGTCTATTGCTTTAATGCAGATCTAAAATCTCCACCAGACTCTGAAGATGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1276 TCACAGGTTGTCTATTGCTTTAATGCAGATCTAAAATCTCCACCAGACTCTGAAGATGGT 1335

Qy       1321 GGGAAAACAATGACTCCAGACTTTACAACAGGAGAAGTAAAATACACTCATATTGCAGGT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1336 GGGAAAACAATGACTCCAGACTTTACAACAGGAGAAGTAAAATACACTCATATTGCAGGT 1395

Qy       1381 CGTGACCTCTTTAAATATACTGTGAAACCAAGAGATACCGATCCTGACACTTTCTTAAAA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1396 CGTGACCTCTTTAAATATACTGTGAAACCAAGAGATACCGATCCTGACACTTTCTTAAAA 1455

Qy       1441 CATATCAAAAAAGTAATTGAGAAGGGTTACAGGGAAAAAGGACAAGCTATTGAGTATAGT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1456 CATATCAAAAAAGTAATTGAGAAGGGTTACAGGGAAAAAGGACAAGCTATTGAGTATAGT 1515

Qy       1501 GGTCTAACTGAGACACAATTGCGTGCGGCTACTCAGTTAGCAATATATTATTTCACTGAT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1516 GGTCTAACTGAGACACAATTGCGTGCGGCTACTCAGTTAGCAATATATTATTTCACTGAT 1575

Qy       1561 AGTGCTGAATTAGATAAGGATAAACTAAAAGACTATCATGGTTTTGGAGACATGAATGAT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1576 AGTGCTGAATTAGATAAGGATAAACTAAAAGACTATCATGGTTTTGGAGACATGAATGAT 1635

Qy       1621 AGTACTTTAGCAGTTGCTAAAATCCTTGTAGAATACGCTCAAGATAGTAATCCTCCACAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1636 AGTACTTTAGCAGTTGCTAAAATCCTTGTAGAATACGCTCAAGATAGTAATCCTCCACAG 1695
```

```
Qy       1681 CTAACTGACCTTGATTTCTTTATTCCGAATAACAATAAATATCAATCTCTTATTGGAACT 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1696 CTAACTGACCTTGATTTCTTTATTCCGAATAACAATAAATATCAATCTCTTATTGGAACT 1755

Qy       1741 CAGTGGCATCCAGAAGATTTAGTTGATATTATTCGTATGGAAGATAAAAAAGAAGTTATA 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1756 CAGTGGCATCCAGAAGATTTAGTTGATATTATTCGTATGGAAGATAAAAAAGAAGTTATA 1815

Qy       1801 CCTGTAACTCATAATTTAACATTGAGAAAAACGGTGACTGGTTTAGCTGGTGACAGAACT 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1816 CCTGTAACTCATAATTTAACATTGAGAAAAACGGTGACTGGTTTAGCTGGTGACAGAACT 1875

Qy       1861 AAAGATTTCCATTTTGAAATTGAATTAAAAAATAATAAGCAAGAATTGCTTTCTCAAACT 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1876 AAAGATTTCCATTTTGAAATTGAATTAAAAAATAATAAGCAAGAATTGCTTTCTCAAACT 1935

Qy       1921 GTTAAAACAGATAAAACAAACCTCGAATTTAAAGATGGTAAAGCAACCATTAATTTAAAA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1936 GTTAAAACAGATAAAACAAACCTCGAATTTAAAGATGGTAAAGCAACCATTAATTTAAAA 1995

Qy       1981 CATGGGGAAAGTTTAACACTTCAAGGTTTACCAGAAGGTTATTCTTACCTTGTCAAAGAA 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1996 CATGGGGAAAGTTTAACACTTCAAGGTTTACCAGAAGGTTATTCTTACCTTGTCAAAGAA 2055

Qy       2041 ACAGATTCTGAAGGCTATAAGGTTAAAGTTAATAGCCAAGAAGTAGCAAATGCTACAGTT 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2056 ACAGATTCTGAAGGCTATAAGGTTAAAGTTAATAGCCAAGAAGTAGCAAATGCTACAGTT 2115

Qy       2101 TCAAAAACAGGAATAACAAGTGATGAGACACTTGCTTTTGAAAATAATAAAGAGCCTGTT 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2116 TCAAAAACAGGAATAACAAGTGATGAGACACTTGCTTTTGAAAATAATAAAGAGCCTGTT 2175

Qy       2161 GTTCCTACAGGAGTTGATCAAAAGATCAATGGCTATCTAGCTTTGATAGTTATCGCTGGT 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2176 GTTCCTACAGGAGTTGATCAAAAGATCAATGGCTATCTAGCTTTGATAGTTATCGCTGGT 2235

Qy       2221 ATCAGTTTGGGGATCTGGGGAATTCACACGATAAGGATAAGAAAACATGAC 2271
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2236 ATCAGTTTGGGGATCTGGGGAATTCACACGATAAGGATAAGAAAACATGAC 2286
```

RESULT 2
ABN69679/c
ID ABN69679 standard; DNA; 159 BP.
XX
AC ABN69679;
XX
DT 01-JUL-2002 (first entry)
XX
DE Streptococcus polynucleotide SEQ ID NO 7271.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX
OS Streptococcus pyogenes.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB04789.
XX
PR 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
PI Tettelin H;
XX
DR WPI; 2002-352536/38.
DR P-PSDB; ABP29048.
XX
PT New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
PS Claim 7; Page 3879; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyrogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
SQ Sequence 159 BP; 46 A; 31 C; 27 G; 55 T; 0 other;

Query Match 3.3%; Score 75; DB 24; Length 159;
Best Local Similarity 100.0%; Pred. No. 3.5e-07;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2200 GCTTTGATAGTTATCGCTGGTATCAGTTTGGGGATCTGGGGAATTCACACGATAAGGATA 2259
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 GCTTTGATAGTTATCGCTGGTATCAGTTTGGGGATCTGGGGAATTCACACGATAAGGATA 100

Qy       2260 AGAAAACATGACTAG 2274
             |||||||||||||||
Db         99 AGAAAACATGACTAG 85
```

RESULT 3
AAA70259
ID AAA70259 standard; DNA; 4677 BP.
XX
AC AAA70259;
XX
DT 07-NOV-2000 (first entry)
XX
DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:392.
XX
KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
KW antimalarial; malaria; protozóacide; infection; insecticide; ds.
XX
OS Plasmodium falciparum.
XX
PN WO200025728-A2.
XX
PD 11-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US26796.
XX
PR 05-NOV-1998; 98US-0107131.
XX
PA (HOFF/) HOFFMAN S.
PA (CARU/) CARUCCI D.
PA (GARD/) GARDNER M.
PA (VENT/) VENTER J C.
XX
PI Hoffman S, Carucci D, Gardner M, Venter JC;
XX
DR WPI; 2000-365347/31.

```
XX
PT   Proteins encoded by chromosome 2 of the human malarial parasite,
PT   Plasmodium falciparum, useful as antimalarial vaccines and in the
PT   diagnosis of P.falciparum infection -
XX
PS   Disclosure; Page 565-566; 577pp; English.
XX
CC   The present invention describes proteins and their fragments (I) encoded
CC   by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC   Also described are: (1) nucleotide sequences (II) encoding (I); and (2)
CC   vaccines against P. falciparum infection comprising (I) or (II).
CC   (I) and (II) are useful for the development of vaccines against
CC   P. falciparum infection. (I) and polyclonal antisera or a monoclonal
CC   antibody raised to immunogens comprising the sequences of (I), are
CC   useful in the detection of infection with P. falciparum. Furthermore,
CC   (I) (especially when they are rifins or secreted or membrane proteins)
CC   can aid the identification of drugs to treat or prevent P. falciparum
CC   infection, or they can be used to identify drug resistance in
CC   P. falciparum. Sequencing of the Plasmodium chromosome 2 and the
CC   subsequent identification of proteins encoded by it will help to expand
CC   our understanding of parasite biology, a process hampered by the
CC   complexity of the parasitic lifecycle, and provide new targets for
CC   vaccine and drug development. Parasite resistance to drugs and mosquito
CC   resistance to insecticides have led to a resurgence of malaria in many
CC   parts of the world, and there is a pressing need for vaccines and new
CC   drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352 represent nucleotide
CC   and protein sequences given in the present invention, but which are not
CC   specifically mentioned within the specification.
XX
SQ   Sequence 4677 BP; 2106 A; 402 C; 966 G; 1203 T; 0 other;

  Query Match             2.5%;  Score 56.6;  DB 21;  Length 4677;
  Best Local Similarity   43.5%;  Pred. No. 0.01;
  Matches  416;  Conservative    0;  Mismatches  529;  Indels   12;  Gaps    3;

Qy       1084 TTTAAGGTTGAAGCTGGCAAAGTGTATACTATTATTGATGGAAAACAGATTGAAAATCCC 1143
Db       3256 TCTGATTTAAAAGATCTTGAAGAAGATATATTAAAAGAAGTAAAAGAAATCAAAGAACTT 3315

Qy       1144 AATAAAGAGATAGTAGAGCCTTACTCAGTAGAAGCATATAATGATTTTGAAGAATTTAGC 1203
Db       3316 GAAAGTGAAATTTTAGAAGATTATAAAGAATTAAAAACTATTGAAACAGATATTTTAGAA 3375

Qy       1204 GTTTTAACTACACAAAACTATGCAAAATTTTATTATGCAAAAAATAAAAATGGAAGTTCA 1263
Db       3376 GAGAAAAAAGAAATAGAAAAAGATCATTTTGAAAAATTCGAAGAAGAAGCTGAAGAAATA 3435

Qy       1264 CAGGTTGTCTATTGCTTTAATGCAGATCTAAAATCTCCACCAGACTCTGAAGATGGTGGG 1323
Db       3436 AAAGATCTTGAAGCAGATATATTAAAAGAAGTATCTTCATTAGAAGTTGAAGAAGAAAAA 3495

Qy       1324 AAAACAATGACTCCAGACTTTACAACAGGAGAAGTAAAATACACTCATATTGCAGGTCGT 1383
Db       3496 AAATTAGAAGAAGTACACGAATTAAAAGAAGAGGTAGAA------CATATAATAAGTGGT 3549

Qy       1384 GACCTCTTTAAATATACTGTGAAACCAAGAGATACCGATCCTGACACTTTCTTAAAACAT 1443
Db       3550 GATGCGCATATAAAAGGTTTGGAAGAAGATGATTTAGAAGAAGTAGATGATTTAAAAGGA 3609

Qy       1444 ATCAAAAAAGTAATTGAGAAGGGTTACAGGGAAAAAGGACAAGCTATTGAGTATAGTGGT 1503
Db       3610 AGTATATTAGACATGTTAAAGGGAGATATGGAATTAGGGGATATGGATAAGGAAAGTTTA 3669

Qy       1504 CTAACTGAGACACAATTGCGTGCGGCTACTCAGTTAGCAATATATTATTTCACTGATAGT 1563
Db       3670 GAAGATGTAACAGCAAAACTTGGAGAAAGAGTTGAATCCTTAAAAGATGTTTTATCTAGT 3729

Qy       1564 GCTGAATTAGATAAGGATAAACTAAAAGACTATCATGGTTTTGGAGACATGAATGATAGT 1623
Db       3730 GC---ATTAGGCATGGATGAAGAACAAATGAAAACAAGAAAAAAAGCTCAAAGACCTAAA 3786

Qy       1624 ACTTTAGCAGTTGCTAAAATCCTTGTAGAATACGCTCAAGATAGTAATCCTCCACAGCTA 1683
```

```
Db       3787 TTGGAAGAAGTATTATTAAAAGAAGAGGTTAAAGAAGAACCAAAGAAAAAAATAACAAAA 3846

Qy       1684 ACTGACCTTGATTTCTTTATTCCGAATAACAATAAATATCAATCTCTTATTGGAACTCAG 1743
Db       3847 AAGAAAGTAAGGTTTGATATTAAGGATAAGGAACCAAAAGATGAAATAGTAGAAGTTGAA 3906

Qy       1744 TGGCATCCAGAAGATTTAGTTGATATTATTCGTATGGAAGATAAAAAAGAAGTTATACCT 1803
Db       3907 ATGAAAGATGAAGATATAGATGAAGATATAGAAGAAGATGTAGAAGAAGATATAGAAGAA 3966

Qy       1804 GTAACTCATAATTTAACATTGAGAAAAACGGTGACTGGTTTAGCTGGTGACAGAACTAAA 1863
Db       3967 GATAAAGTTGAAGATATAGATGAAGATATAGATGAAGATATAGATGAAGATATAGGTGAA 4026

Qy       1864 GATTTCCATTTTGAAATTGAATTAAAAAATAATAAGCAAGAATTGCTTTCTCAAACTGTT 1923
Db       4027 GACAAAGATGAAGTTATAGATTTAATAGTCCAAAAAGAGAAACGCATTG---AAAAGGTT 4083

Qy       1924 AAAACAGATAAAACAAACCTCGAATTTAAAGATGGTAAAGCAACCATTAATTTAAAACAT 1983
Db       4084 AAAGAGAAAAAGAAAAAATTAGAAAAAAAAGTTGAAGAAGGTGTTAGTGGTCTTAAAAAA 4143

Qy       1984 GGGGAAAGTTTAACACTTCAAGGTTTACCAGAAGGTTATTCTTACCTTGTCAAAGAA 2040
Db       4144 CACGTAGACGAAGTAATGAAATATGTTCAAAAAATTGATAAAGAAGTTGATAAAGAA 4200


RESULT 4
ABL56203
ID   ABL56203 standard; DNA; 32392 BP.
XX
AC   ABL56203;
XX
DT   01-JUL-2002  (first entry)
XX
DE   AmEPV genome fragment#5.
XX
KW   AmEPV; gene therapy; viral vector; chromosome mapping; gene mapping;
KW   genetic deficiency disorder; ds.
XX
OS   Amsacta moorei entomopoxvirus.
XX
PN   WO200212526-A2.
XX
PD   14-FEB-2002.
XX
PF   10-AUG-2001; 2001WO-US25287.
XX
PR   10-AUG-2000; 2000US-224479P.
PR   14-SEP-2000; 2000US-0662254.
XX
PA   (UYFL ) UNIV FLORIDA.
XX
PI   Moyer RW,  Li Y,  Bawden AL;
XX
DR   WPI; 2002-227161/28.
XX
PT   Novel recombinant entomopox virus vector useful for delivering
PT   polynucleotide encoding protein to vertebrate cell, comprises
PT   polynucleotide encoding protein operably linked with heterologous
PT   promoter sequence  -
XX
PS   Disclosure; Page 226-242; 326pp; English.
XX
CC   The invention relates to a recombinant entomopox virus (EPV) vector,
CC   comprising a polynucleotide encoding a protein operably linked with a
CC   heterologous promoter sequence. The invention also concerns methods for
CC   providing gene therapy for genetic deficiency disorders. Vectors of the
CC   invention are useful for delivering a polynucleotide encoding a protein
CC   to a vertebrate cell preferably a mammalian cell, such as a human cell.
CC   The vector is introduced into the vertebrate cell by infection in a viral
CC   particle, or by transfection, transduction, or injection either in vitro
CC   or in vivo. The vector is useful for the delivery and expression of
```

```
CC   biologically useful proteins in gene therapy protocols, and for
CC   delivering large DNA segments for engineering of vertebrate cells.
CC   Polynucleotides of the invention have applications in techniques such as
CC   their use as insertion sites for foreign genes of interest, hybridisation
CC   probes, for chromosome and gene mapping, in PCR technologies, and in the
CC   production of sense or antisense nucleic acids. Vectors of the invention
CC   provide for stable integration and expression of heterologous DNA in host
CC   cells, and are adapted for accepting large heterologous polynucleotide
CC   inserts which can be delivered in an infected or transformed cell and
CC   expressed in a stable fraction. The current sequence represents a
CC   fragment of the genome of the genus B entomopoxvirus from amsacta moorei
CC   (AmEPV).
XX
SQ   Sequence 32392 BP; 13748 A; 2577 C; 2550 G; 13517 T; 0 other;

  Query Match             2.5%;  Score 56.6;  DB 24;  Length 32392;
  Best Local Similarity   44.0%;  Pred. No. 0.017;
  Matches  427;  Conservative    0;  Mismatches  534;  Indels   10;  Gaps    4;

Qy       1065 TATCGCAGAGCCAATCACTTTTAAGGTTGAAGCTGGCAAAGTGTATACTATTATTGATGG 1124
Db       7205 TCTTGAAGTAAAAAATTCATATATTGAAGATGATAAAATGTTATTATATGTTAATAAAAA 7264

Qy       1125 AAAACAGATTGAAAATCCCAATAAAGAGATAGTAGAGCCTTACTCAGTAGAAGCATATAA 1184
Db       7265 TTTACAATTTTTAAAAATATATGAATCTAAATTAGAATTATCCTATATTGAAGAATTTAT 7324

Qy       1185 TGATTTTGAAGAATTTAGCGTTTTAACTACACAAAACTATGCAAAATTTTATTATGCAAA 1244
Db       7325 AAATTTAAAATATATAAGTGTGTGGGACGAAAAAACTTATATAAATAGTAAATATTTAAA 7384

Qy       1245 AAATAAAAATGGAAGTTCACAGGTTGTCTATTGCTTTAATGCAGATCTAAAATCTCCACC 1304
Db       7385 AAATATGTATAATTTAAACTATATTGAAATTTTTAATATTATTAATATTAATGGTTTGAT 7444

Qy       1305 AGAC---TCTGAAGATGGTGGGAAAACAATGACTCCAGACTTTACAACAGGAGAAGTAAA 1361
Db       7445 AAACATATATGATTTAAAATTTTTAAGAACTAATTTAAATAATAAATTTGTTGATATAAA 7504

Qy       1362 ATACACTCATATTGCAGGTCGTGAC--CTCTTTAAATATACTGTGAAACCAAGAGATACC 1419
Db       7505 ATTATTAAATTTTCTACACAATACCGAATGTTTAGACATAATATGTAATAAAAATACAAT 7564

Qy       1420 GATCCTGACACTT---TCTTAAAACATATCAAAAAAGTAATTGAGAAGGGTTACAGGGAA 1476
Db       7565 TATAAAATCATTTGAATATTTAACTAAATTAATAAAATTATCTATATATTATTACAATAA 7624

Qy       1477 AAAGGACAAGCTATTGAGTATAGTGGTCTAACTGAGACACAATTGCGTGCGGCTACTCAG 1536
Db       7625 AATTAATAGTATAATGTATAATTTAAATTCATCGAATTTACAAATATTAAATATTTGTAC 7684

Qy       1537 TTAGCAATATATTATTTCACTGATAGTGCTGAATTAGATAAGGATAAACTAAAAGACTAT 1596
Db       7685 TAATAATTATATTGATTTTAAATTATTTAAAAACTTGTTAAATATAAAATATATGAAATT 7744

Qy       1597 CATGGTTTTGGAGACATGAATGATAGTACTTTAGCAGTTGCTAAAATCCTTGTAGAATAC 1656
Db       7745 AATAAATATAAATAAAAAAAATAAAATACGTTTTGATGTAAATAATATATTAAAGTTTAA 7804

Qy       1657 GCTCAAGATAGTAATCCTCCACAGCTAACTGACCTTGATTTCTTTATTCCGAATAACAAT 1716
Db       7805 AAGTTTATTATCATT--AAAAATTGAAAATATGCATATAGATAATATTGAAAAAATCAGT 7862

Qy       1717 AAATATCAATCTCTTATTGGAACTCAGTGGCATCCAGAAGATTTAGTTGATATTATTCGT 1776
Db       7863 AATTTTAATACTATTGAAGTATTACACTTAAATAATATTGATATAGTAAATATAAATTTT 7922

Qy       1777 ATGGAAGATAAAAAAGAAGTTATACCTGTAACTCATAATTTAACATTGAGAAAAACGGTG 1836
Db       7923 ATAGAAAATAATTTAAATTTAATCGAATTAAATTTAGATAATAATTATATAAATAATATA 7982

Qy       1837 ACTGGTTTAGCTGGTGACAGAACTAAAGATTTCCATTTTGAAATTGAATTAAAAAATAAT 1896
```

```
Db       7983 AATTCTTTAAAATGTTTAAAAAAAATAAAAAAATTATCATTAAAAACAAACAATATTATT 8042

Qy       1897 AAGCAAGAATTGCTTTCTCAAACTGTTAAAACAGATAAAACAAACCTCGAATTTAAAGAT 1956
Db       8043 GATATAAAACCATTATTATACTTAAATAATTTAAATTATATTAATATAAAATATAATAAT 8102

Qy       1957 GGTAAAGCAACCATTAATTTAAAACATGGGGAAAGTTTAACACTTCAAGGTTTACCAGAA 2016
Db       8103 ATAAAAGATATTAATATTTTAAGATATTTTAAAAATAATATTGATTTATATATAGATGAT 8162

Qy       2017 GGTTATTCTTA 2027
Db       8163 TATTATATTCA 8173
```

```
RESULT 5
ABQ67060/c
ID   ABQ67060 standard; DNA; 34688 BP.
XX
AC   ABQ67060;
XX
DT   28-AUG-2002  (first entry)
XX
DE   Human angiogenesis associated polynucleotide SEQ ID NO 90.
XX
KW   Human; angiogenesis; methylation; eye disease; glaucoma; tumour;
KW   inflammation; rheumatoid arthritis; diabetic retinopathy; antiulcerds;
KW   macular degeneration; inflammatory bowel disease; Crohn's disease;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiarteriosclerotic; ds.
XX
OS   Homo sapiens.
XX
PN   WO200246454-A2.
XX
PD   13-JUN-2002.
XX
PF   06-DEC-2001; 2001WO-EP14320.
XX
PR   06-DEC-2000; 2000DE-1061338.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Schacht O;
XX
DR   WPI; 2002-500450/53.
XX
PT   New nucleic acid fragments from chemically treated
PT   angiogenesis-associated genes, useful for determining methylation
PT   status, e.g. in diagnosis or treatment of cancer  -
XX
PS   Claim 1; SEQ ID NO 90; 41pp + Sequence Listing; German.
XX
CC   The invention relates to a nucleic acid (I) comprising a segment of 18
CC   bases of chemically pretreated DNA of angiogenesis-associated genes (II)
CC   having sequences (ABQ66971-ABQ67178) or their complements. (I), also
CC   related oligomers, are used to evaluate the methylation status and/or
CC   single-nucleotide polymorphisms, in angiogenesis-related genes, for
CC   diagnosis and treatment of eye diseases, proliferative retinopathy,
CC   neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,
CC   diabetic retinopathy, macular degeneration caused by neovascularisation,
CC   psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and
CC   Crohn's disease.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 34688 BP; 11855 A; 444 C; 5972 G; 16417 T; 0 other;

  Query Match             2.5%;  Score 56.2;  DB 24;  Length 34688;
  Best Local Similarity   51.8%;  Pred. No. 0.022;
  Matches  127;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;
```

```
Qy       1871 ATTTTGAAATTGAATTAAAAAATAATAAGCAAGAATTGCTTTCTCAAACTGTTAAAACAG 1930
Db      26859 AATATAAAACTTACTAAAAAAATAACTTCCCAAAATTTAATTTTCTATTAAACCAAAAAT 26800

Qy       1931 ATAAAACAAACCTCGAATTTAAAGATGGTAAAGCAACCATTAATTTAAAACATGGGGAAA 1990
Db      26799 ATTATTTTAAAAAAATATTTACAAATTTTAAAATTAACATTAATTAAAAATATTACTAAA 26740

Qy       1991 GTTTAACACTTCAAGGTTTACCAGAAGGTTATTCTTACCTTGTCAAAGAAACAGATTCTG 2050
Db      26739 ACTAAACTAAACCCGCTTACCCAAAAAATAATTATAAAAAAATTCTTTTATAACTTCCAA 26680

Qy       2051 AAGGCTATAAGGTTAAAGTTAATAGCCAAGAAGTAGCAAATGCTACAGTTTCAAAAACAG 2110
Db      26679 AACAATCTACTATTAACATAAATTACTAACAATCATCCAAATATATAAAAACAAAAAATT 26620

Qy       2111 GAATA 2115
Db      26619 AAATA 26615
```

RESULT 6
AAA70187
ID AAA70187 standard; DNA; 11922 BP.
XX
AC AAA70187;
XX
DT 07-NOV-2000 (first entry)
XX
DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:320.
XX
KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
KW antimalarial; malaria; protozoacide; infection; insecticide; ds.
XX
OS Plasmodium falciparum.
XX
PN WO200025728-A2.
XX
PD 11-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US26796.
XX
PR 05-NOV-1998; 98US-0107131.
XX
PA (HOFF/) HOFFMAN S.
PA (CARU/) CARUCCI D.
PA (GARD/) GARDNER M.
PA (VENT/) VENTER J C.
XX
PI Hoffman S, Carucci D, Gardner M, Venter JC;
XX
DR WPI; 2000-365347/31.
XX
PT Proteins encoded by chromosome 2 of the human malarial parasite,
PT Plasmodium falciparum, useful as antimalarial vaccines and in the
PT diagnosis of P.falciparum infection -
XX
PS Disclosure; Page 516-519; 577pp; English.
XX
CC The present invention describes proteins and their fragments (I) encoded
CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC Also described are: (1) nucleotide sequences (II) encoding (I); and (2)
CC vaccines against P. falciparum infection comprising (I) or (II).
CC (I) and (II) are useful for the development of vaccines against
CC P. falciparum infection. (I) and polyclonal antisera or a monoclonal
CC antibody raised to immunogens comprising the sequences of (I), are
CC useful in the detection of infection with P. falciparum. Furthermore,
CC (I) (especially when they are rifins or secreted or membrane proteins)
CC can aid the identification of drugs to treat or prevent P. falciparum
CC infection, or they can be used to identify drug resistance in
CC P. falciparum. Sequencing of the Plasmodium chromosome 2 and the
CC subsequent identification of proteins encoded by it will help to expand
CC our understanding of parasite biology, a process hampered by the
CC complexity of the parasitic lifecycle, and provide new targets for
CC vaccine and drug development. Parasite resistance to drugs and mosquito
CC resistance to insecticides have led to a resurgence of malaria in many
CC parts of the world, and there is a pressing need for vaccines and new
CC drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352 represent nucleotide
CC and protein sequences given in the present invention, but which are not
CC specifically mentioned within the specification.
XX
SQ Sequence 11922 BP; 5402 A; 948 C; 1343 G; 4229 T; 0 other;

Query Match 2.5%; Score 56; DB 21; Length 11922;
Best Local Similarity 41.3%; Pred. No. 0.018;
Matches 380; Conservative 0; Mismatches 540; Indels 0; Gaps 0;

```
Qy        988 AATGATATTGGAGAAAGAATTGAACTATCAGATGGAACTTATACTTTAACTGAATTGAAT 1047
Db      10261 AATACAAATACAAATAAAAATTCTGATACAGATAATCATAATGATGCATATAATCATAAT 10320

Qy       1048 TCTCCAGCTGGTTATAGTATCGCAGAGCCAATCACTTTTAAGGTTGAAGCTGGCAAAGTG 1107
Db      10321 GATGCATATAATCATAATGATGCATATAATCATAATGACGCATATAATCATAATGATGCA 10380

Qy       1108 TATACTATTATTGATGGAAAACAGATTGAAAATCCCAATAAAGAGATAGTAGAGCCTTAC 1167
Db      10381 TATAATCATAATGATGCATATAATCATAATGATGCATATAATCATAATGATGCATATAAT 10440

Qy       1168 TCAGTAGAAGCATATAATGATTTTGAAGAATTTAGCGTTTTAACTACACAAAACTATGCA 1227
Db      10441 CATAATGATGCATATAATCATAATGATGCAGATAATCATAATGATACAGATAATCATAGT 10500

Qy       1228 AAATTTTATTATGCAAAAAATAAAAATGGAAGTTCACAGGTTGTCTATTGCTTTAATGCA 1287
Db      10501 GATAATTATAATTCTCATAAATATAAAGGAACTTATAAAATTTATCGTATTCATGATGAG 10560

Qy       1288 GATCTAAAATCTCCACCAGACTCTGAAGATGGTGGGAAAACAATGACTCCAGACTTTACA 1347
Db      10561 GAAGATATTATTCAAGATAATAATTATACAAACGATGATTATGTTAATACAAATTATAAT 10620

Qy       1348 ACAGGAGAAGTAAAATACACTCATATTGCAGGTCGTGACCTCTTTAAATATACTGTGAAA 1407
Db      10621 TTTCATCAAAATAATAATTATCAAAACAGCTCCAAGGATAATAGCCATACTTATACTAGT 10680

Qy       1408 CCAAGAGATACCGATCCTGACACTTTCTTAAAACATATCAAAAAAGTAATTGAGAAGGGT 1467
Db      10681 TCGGGTGATGATAATTTAAATAATCCATTTATATCTTTAAATAAAGAAAATATATCCAAA 10740

Qy       1468 TACAGGGAAAAAGGACAAGCTATTGAGTATAGTGGTCTAACTGAGACACAATTGCGTGCG 1527
Db      10741 AAAAAAAAAAACAAAAAAACAAATAAAAATAGAGAAGAAGAAAAAAAAAAGAAAAATGAA 10800

Qy       1528 GCTACTCAGTTAGCAATATATTATTTCACTGATAGTGCTGAATTAGATAAGGATAAACTA 1587
Db      10801 ATTATACATAATGAAGTAGAATATTTCTTAGAAGCTATTTTAAATGTTGTAGGGAAATTA 10860

Qy       1588 AAAGACTATCATGGTTTTGGAGACATGAATGATAGTACTTTAGCAGTTGCTAAAATCCTT 1647
Db      10861 TCAGATCGTATTAATTTTATATTTTTAAAAATCGAAAAAATTAATTGTCTAGCTATTTTT 10920

Qy       1648 GTAGAATACGCTCAAGATAGTAATCCTCCACAGCTAACTGACCTTGATTTCTTTATTCCG 1707
Db      10921 GAAGAATCCTATCTATATGTTGAACTACTAACAAAAATATCTTATTATTGTAATATTTAT 10980

Qy       1708 AATAACAATAAATATCAATCTCTTATTGGAACTCAGTGGCATCCAGAAGATTTAGTTGAT 1767
Db      10981 AATATAAATAATTTCTTAATATCTTTAATGGATAAAAGAGTTGAACATCATTTATTTTAT 11040

Qy       1768 ATTATTCGTATGGAAGATAAAAAAGAAGTTATACCTGTAACTCATAATTTAACATTGAGA 1827
Db      11041 ATAAATAGAATATTAGAAAAATTCATTTACGAAAAAGAAGTTATTATATATCCATATAGT 11100

Qy       1828 AAAACGGTGACTGGTTTAGCTGGTGACAGAACTAAAGATTTCCATTTTGAAATTGAATTA 1887
Db      11101 CCTTATGAAATTAATGCATCAACAAATGTACCTAATAATCAAGATAAAAAAAAAAAAAAA 11160
```

```
Qy       1888 AAAAATAATAAGCAAGAATT 1907
              |||   |  ||  | ||| |
Db      11161 AAATCAACAAAAAATGAACT 11180
```

RESULT 7
AAA70105
ID AAA70105 standard; DNA; 5940 BP.
XX
AC AAA70105;
XX
DT 07-NOV-2000 (first entry)
XX
DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:238.
XX
KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
KW antimalarial; malaria; protozoacide; infection; insecticide; ds.
XX
OS Plasmodium falciparum.
XX
PN WO200025728-A2.
XX
PD 11-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US26796.
XX
PR 05-NOV-1998; 98US-0107131.
XX
PA (HOFF/) HOFFMAN S.
PA (CARU/) CARUCCI D.
PA (GARD/) GARDNER M.
PA (VENT/) VENTER J C.
XX
PI Hoffman S, Carucci D, Gardner M, Venter JC;
XX
DR WPI; 2000-365347/31.
XX
PT Proteins encoded by chromosome 2 of the human malarial parasite,
PT Plasmodium falciparum, useful as antimalarial vaccines and in the
PT diagnosis of P.falciparum infection -
XX
PS Disclosure; Page 460-462; 577pp; English.
XX
CC The present invention describes proteins and their fragments (I) encoded
CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC Also described are: (1) nucleotide sequences (II) encoding (I); and (2)
CC vaccines against P. falciparum infection comprising (I) or (II).
CC (I) and (II) are useful for the development of vaccines against
CC P. falciparum infection. (I) and polyclonal antisera or a monoclonal
CC antibody raised to immunogens comprising the sequences of (I), are
CC useful in the detection of infection with P. falciparum. Furthermore,
CC (I) (especially when they are rifins or secreted or membrane proteins)
CC can aid the identification of drugs to treat or prevent P. falciparum
CC infection, or they can be used to identify drug resistance in
CC P. falciparum. Sequencing of the Plasmodium chromosome 2 and the
CC subsequent identification of proteins encoded by it will help to expand
CC our understanding of parasite biology, a process hampered by the
CC complexity of the parasitic lifecycle, and provide new targets for
CC vaccine and drug development. Parasite resistance to drugs and mosquito
CC resistance to insecticides have led to a resurgence of malaria in many
CC parts of the world, and there is a pressing need for vaccines and new
CC drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352 represent nucleotide
CC and protein sequences given in the present invention, but which are not
CC specifically mentioned within the specification.
XX
SQ Sequence 5940 BP; 3106 A; 343 C; 879 G; 1612 T; 0 other;

Query Match 2.4%; Score 53.8; DB 21; Length 5940;
Best Local Similarity 44.7%; Pred. No. 0.046;
Matches 294; Conservative 0; Mismatches 357; Indels 6; Gaps 2;

```
Qy       1477 AAAGGACAAGCTATTGAGTATAGTGGTCTAACTGAGACACAATTGCGTGCGGCTACTCAG 1536
              || |   ||| || |||  ||| ||     |  || | |           |  ||   |
Db         73 AATGAGGAAGATAATGAAAATAATGAAAAGAGGGAAAAAAGCGATTCATGGTATAAAAAA 132

Qy       1537 TTAGCAATATATTATTTCACTGATAGTGCTGAATTAGATAAGGATAAACTAAAAGACTAT 1596
              ||  |  |  | |    |    ||   || |||   | || |||||      ||| ||
Db        133 ATAATAGAAACTAAAGGAAAAAGTAAAACTAAATATAAAAATGATAATTCCTTAGATGAT 192

Qy       1597 CATGGTTTTGGAGACAT---GAATGATAGTACTTTAGCAGTTGCTAAAATCCTTGTAGAA 1653
              ||  |  ||  |||||    ||| ||| || |         |  | | |    |   ||
Db        193 AATATTAATGAGGACATAATAAATAATAATAATAATAATAATAATGATAATAATAATGAT 252

Qy       1654 TACGCTCAAGATAGTAATCCTCCACAGCTAACTGACCTTGATTTCTTTATTCCGAATAAC 1713
              ||  | | |||| ||||  |    |    | |||   | ||     || |   ||| |
Db        253 AACAATAATGATAATAATAATGATAATAATAATGATAATAATAATGATAATAATAATGAG 312

Qy       1714 AATAAATATCAATCTCTTATTGGAACTCAGTGGCATCCAGAAGATTTAGTTGATATTATT 1773
              |||||  || |   |  || |   | | | |         |||   | |   |||||||
Db        313 AATAATAATGATAATAATAATTTTAATAATTATAGCGATGAAATATCAAAAAATATTATA 372

Qy       1774 CGTATGGAAGATAAAAAAGAAGTTATACCTGTAACTCATAATTTAACATTGAGAAAAACG 1833
              | ||  ||  || |   ||||      | |      |          |  | ||    |
Db        373 CATAAAGACAATGAGCTAGAAAACCAGCTTAAGGATACATTAAAGTCCATTAGTTCGTTG 432

Qy       1834 GTGACTGGTTTAGCTGGTGACAGAACTAAAGATTTCCATTTTGAAATTGAATTAAAAAAT 1893
              || |    | |    | ||  || ||||  |    | || ||||  |||||||| |
Db        433 TCGAATAAAATTGTGAATTACGAAAGTAAAATTGAAGAATTAGAAAAAGAATTAAAAGAA 492

Qy       1894 AATAAGCAAGAATTGCTTTCTCAAACTGTTAAAACAGATAAAACAAACCTCGAATTTAAA 1953
              ||| |  |     ||  | | | || | |   | |||||  |||     ||   |
Db        493 GTAAAGGATAAGAATATTGATAATAATGATTATGAAAATAAATTAAAAGAAAAAGAAGAT 552

Qy       1954 GATGGTAAAGCAACCATTAATTTAAAACATGGGGAAAGTTTAACACTTCAAGGTTTACCA 2013
              ||  ||||  ||  ||| || |   | ||   ||||    ||  |||  |    |  |
Db        553 TTTGTTAAACAAAAAATTGATATGCTAAAT---GAAAAAGAAAATCTTTTACAAGAAAAA 609

Qy       2014 GAAGGTTATTCTTACCTTGTCAAAGAAACAGATTCTGAAGGCTATAAGGTTAAAGTTAAT 2073
              |||    ||  | |        || | | |  |  ||||    | |||  || | | ||
Db        610 GAATTAGATATTAATAAAAGAGAAAAGAAAATTAATGAAAAAGAAAAGAATATAATAAAA 669

Qy       2074 AGCCAAGAAGTAGCAAATGCTACAGTTTCAAAAACAGGAATAACAAGTGATGAGACA 2130
              |  |||||  |    ||  || ||    | |      |  || || | | || | |
Db        670 AAGGAAGAAACATTTCATAATATAGAAAAAGAGTATTTAGAAAAAAATAAAGAAAGA 726
```

RESULT 8
ABV10021
ID ABV10021 standard; cDNA; 494 BP.
XX
AC ABV10021;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 10012.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

```
XX
PI   Schlegel R,  Endege WO,  Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS   Claim 1; Page 1602; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement. (I) is useful for:
CC   (a) assessing whether a patient is afflicted with prostate cancer;
CC   (b) monitoring the progression of prostate cancer in a patient;
CC   (c) assessing the efficacy of a test compound to inhibit prostate
CC   cancer in a patient;
CC   (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC   in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound;
CC   (g) determining whether prostate cancer has metastasized in a patient;
CC   (h) assessing the aggressiveness or indolence of prostate cancer in a
CC   patient;
CC   (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
XX
SQ   Sequence 494 BP; 299 A; 19 C; 21 G; 155 T; 0 other;

  Query Match             2.3%;  Score 53;  DB 23;  Length 494;
  Best Local Similarity   48.2%;  Pred. No. 0.036;
  Matches  149;  Conservative    0;  Mismatches  160;  Indels    0;  Gaps    0;

Qy       1766 ATATTATTCGTATGGAAGATAAAAAAGAAGTTATACCTGTAACTCATAATTTAACATTGA 1825
Db        180 AAATTAAAAAAATAAAAAAAAAAAAAAAAAATTAAAAATATAAAAAAAAAAAAAAAAAT 239

Qy       1826 GAAAAACGGTGACTGGTTTAGCTGGTGACAGAACTAAAGATTTCCATTTTGAAATTGAAT 1885
Db        240 AAAAAATAAAAAAATAAATATAAATAAAAAAATAAAAAAAAAAATATAAAATAAAAAAAA 299

Qy       1886 TAAAAAATAATAAGCAAGAATTGCTTTCTCAAACTGTTAAAACAGATAAAACAAACCTCG 1945
Db        300 AAAAAAAAAAAAATATTATAAAAATATATAAAAAAAATAAAATAAATATAAAAAAATTAA 359

Qy       1946 AATTTAAAGATGGTAAAGCAACCATTAATTTAAAACATGGGGAAAGTTTAACACTTCAAG 2005
Db        360 TATAAATAAAATATAAAAAAATAATATAATAAAAAAATTAAAAAAAATAATACAATAAAA 419

Qy       2006 GTTTACCAGAAGGTTATTCTTACCTTGTCAAAGAAACAGATTCTGAAGGCTATAAGGTTA 2065
Db        420 TAAAAAACAAAAATAAATAATATTTTAAAAAATATATAAATTATAAAATATATAAATATT 479

Qy       2066 AAGTTAATA 2074
Db        480 AATATAATA 488


RESULT 9
AAI61371
ID   AAI61371 standard; DNA; 335913 BP.
XX
AC   AAI61371;
XX
DT   16-OCT-2001  (first entry)
XX
DE   Soybean 240017 region G3, SEQ ID NO: 2.
XX
KW   Soybean; antihelmintic; gene therapy; soybean cyst nematode; SCN;
KW   SCN resistance; rhg1; Rhg4; SCN resistant allele; plant breeding;
KW   240017 region G3; 318013 region A3; 515002 region G2; ds.
XX
OS   Glycine max.
XX
PN   WO200151627-A2.
XX
PD   19-JUL-2001.
XX
PF   05-JAN-2001; 2001WO-US00552.
XX
PR   07-JAN-2000; 2000US-0174880.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Hauge BM,  Wang ML,  Parsons JD,  Parnell LD;
XX
DR   WPI; 2001-425872/45.
DR   P-PSDB; AAM42214.
XX
PT   New purified nucleic acid for producing a soybean plant having soybean
PT   cyst nematode resistance and for use in plant breeding programs -
XX
PS   Claim 2; Page 204-400; 1353pp; English.
XX
CC   The invention relates to nucleic acid molecules from regions of the
CC   soybean genome which are associated with soybean cyst nematode (SCN)
CC   resistance. The nucleic acids are used to transform plants, and can
CC   produce soybean plants having an rhg1 or an Rhg4 SCN resistant allele.
CC   The nucleic acids can be used for investigating rhg1 or Rhg4 haplotypes
CC   of soybean plants and for introgressing SCN resistance or partial SCN
CC   resistance into soybean plants. They can also be used in plant breeding
CC   programmes. The invention also relates to proteins encoded by such
CC   nucleic acid molecules, as well as antibodies capable of recognising
CC   these proteins. The present sequence is a nucleic acid molecule
CC   provided in the specification.
XX
SQ   Sequence 335913 BP; 114579 A; 53403 C; 53026 G; 114905 T; 0 other;

  Query Match             2.3%;  Score 53;  DB 22;  Length 335913;
  Best Local Similarity   45.7%;  Pred. No. 0.2;
  Matches  185;  Conservative    0;  Mismatches  220;  Indels    0;  Gaps    0;

Qy       1712 ACAATAAATATCAATCTCTTATTGGAACTCAGTGGCATCCAGAAGATTTAGTTGATATTA 1771
Db      99367 ACAAAATATTAAAATATATAAAATTATCTCAATTTTAAATATACAAAATATTTTAAATAA 99426

Qy       1772 TTCGTATGGAAGATAAAAAAGAAGTTATACCTGTAACTCATAATTTAACATTGAGAAAAA 1831
Db      99427 AACATTATTAAAAAATAATATAAAATATTTAAAATAAAAAATATATAACATATTAAATAA 99486

Qy       1832 CGGTGACTGGTTTAGCTGGTGACAGAACTAAAGATTTCCATTTTGAAATTGAATTAAAAA 1891
Db      99487 AATCATATAAAAAATACATACAAAATATTAAATAAAACTAATAAGAAGTAAAATTAAAAA 99546

Qy       1892 ATAATAAGCAAGAATTGCTTTCTCAAACTGTTAAAACAGATAAAACAAACCTCGAATTTA 1951
Db      99547 TATTTATTTAATAATTAAATAAATAAAATTATTTACAACTTTTTAAAAAATTTGAAAATA 99606

Qy       1952 AAGATGGTAAAGCAACCATTAATTTAAAACATGGGGAAAGTTTAACACTTCAAGGTTTAC 2011
Db      99607 AAAACAAAAAAGAAAATACAAGTAAATAAAAATCCTAAAACAAAAAACTTTAAACATGAA 99666

Qy       2012 CAGAAGGTTATTCTTACCTTGTCAAAGAAACAGATTCTGAAGGCTATAAGGTTAAAGTTA 2071
Db      99667 AAAAAATGACTTTAAAGTCATAGAAAAAAAAAAACTTACAACTTCAAAGTGGTAAGTTCA 99726

Qy       2072 ATAGCCAAGAAGTAGCAAATGCTACAGTTTCAAAAACAGGAATAA 2116
Db      99727 AAGAAAAAACTTGTACGACTTTAAAAATGTAAATAAAATAAAAAA 99771


RESULT 10
AAI61372
ID   AAI61372 standard; DNA; 335913 BP.
XX
AC   AAI61372;
```

```
XX
DT   16-OCT-2001  (first entry)
XX
DE   Soybean 240017 region G3, SEQ ID NO: 3.
XX
KW   Soybean; antihelmintic; gene therapy; soybean cyst nematode; SCN;
KW   SCN resistance; rhg1; Rhg4; SCN resistant allele; plant breeding;
KW   240017 region G3; 318013 region A3; 515002 region G2; ds.
XX
OS   Glycine max.
XX
PN   WO200151627-A2.
XX
PD   19-JUL-2001.
XX
PF   05-JAN-2001; 2001WO-US00552.
XX
PR   07-JAN-2000; 2000US-0174880.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Hauge BM,  Wang ML,  Parsons JD,  Parnell LD;
XX
DR   WPI; 2001-425872/45.
DR   P-PSDB; AAM42215.
XX
PT   New purified nucleic acid for producing a soybean plant having soybean
PT   cyst nematode resistance and for use in plant breeding programs -
XX
PS   Claim 2; Page 400-595; 1353pp; English.
XX
CC   The invention relates to nucleic acid molecules from regions of the
CC   soybean genome which are associated with soybean cyst nematode (SCN)
CC   resistance. The nucleic acids are used to transform plants, and can
CC   produce soybean plants having an rhg1 or an Rhg4 SCN resistant allele.
CC   The nucleic acids can be used for investigating rhg1 or Rhg4 haplotypes
CC   of soybean plants and for introgressing SCN resistance or partial SCN
CC   resistance into soybean plants. They can also be used in plant breeding
CC   programmes. The invention also relates to proteins encoded by such
CC   nucleic acid molecules, as well as antibodies capable of recognising
CC   these proteins. The present sequence is a nucleic acid molecule
CC   provided in the specification.
XX
SQ   Sequence 335913 BP; 114582 A; 53398 C; 53027 G; 114906 T; 0 other;

  Query Match             2.3%;  Score 53;  DB 22;  Length 335913;
  Best Local Similarity   45.7%;  Pred. No. 0.2;
  Matches  185;  Conservative    0;  Mismatches  220;  Indels    0;  Gaps    0;

Qy       1712 ACAATAAATATCAATCTCTTATTGGAACTCAGTGGCATCCAGAAGATTTAGTTGATATTA 1771
             |||| | ||   ||| | | |    | |||| |   |   | |  |  || || | || |
Db      99367 ACAAAATATTAAAATATATAAAATTATCTCAATTTTAAATATACAAAATATTTTAAATAA 99426

Qy       1772 TTCGTATGGAAGATAAAAAAGAAGTTATACCTGTAACTCATAATTTAACATTGAGAAAAA 1831
               | |    || | | || | ||  |||       |   |  || ||||||    || ||
Db      99427 AACATTATTAAAAAATAATATAAAATATTTAAAATAAAAAATATATAACATATTAAATAA 99486

Qy       1832 CGGTGACTGGTTTAGCTGGTGACAGAACTAAAGATTTCCATTTTGAAATTGAATTAAAAA 1891
                  |     |       | |  | |||| |   | | |  ||| |  |||||||||
Db      99487 AATCATATAAAAAATACATACAAAATATTAAATAAAACTAATAAGAAGTAAAATTAAAAA 99546

Qy       1892 ATAATAAGCAAGAATTGCTTTCTCAAACTGTTAAAACAGATAAAACAAACCTCGAATTTA 1951
                 ||   || ||||   |    ||| |  | |  |  |  |   | |||  | |||  ||
Db      99547 TATTTATTTAATAATTAAATAAATAAAATTATTTACAACTTTTTAAAAAATTTGAAAATA 99606

Qy       1952 AAGATGGTAAAGCAACCATTAATTTAAAACATGGGGAAAGTTTAACACTTCAAGGTTTAC 2011
               || |    |||| ||  |  | |  | || |     |||    || |||| ||   | |
Db      99607 AAAACAAAAAAGAAAATACAAGTAAATAAAAATCCTAAAACAAAAAACTTTAAACATGAA 99666

Qy       2012 CAGAAGGTTATTCTTACCTTGTCAAAGAAACAGATTCTGAAGGCTATAAGGTTAAAGTTA 2071
               | ||     ||   |       ||| ||| | | |   ||    | |  | ||| | |
Db      99667 AAAAAATGACTTTAAAGTCATAGAAAAAAAAAAACTTACAACTTCAAAGTGGTAAGTTCA 99726

Qy       2072 ATAGCCAAGAAGTAGCAAATGCTACAGTTTCAAAAACAGGAATAA 2116
                 |     ||       | | |   | | | | || || |  || ||
Db      99727 AAGAAAAAACTTGTACGACTTTAAAAATGTAAATAAAATAAAAAA 99771


RESULT 11
ABL70331/c
ID   ABL70331 standard; DNA; 8210 BP.
XX
AC   ABL70331;
XX
DT   01-JUL-2002  (first entry)
XX
DE   Chemically treated cell signalling DNA sequence#111.
XX
KW   Cell signalling; cytosine methylation; cell signalling disease;
KW   cancer; tumour; cytostatic; ds.
XX
OS   Unidentified.
XX
PN   WO200202807-A2.
XX
PD   10-JAN-2002.
XX
PF   29-JUN-2001; 2001WO-EP07471.
XX
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-154758/20.
XX
PT   Nucleic acid, useful for diagnosis and therapy of diseases associated
PT   with cell signalling e.g. cancer, comprises chemically modified genomic
PT   sequences of genes associated with cell signalling  -
XX
PS   Claim 1; SEQ ID NO 221; 24pp+sequence listing; English.
XX
CC   The invention relates to a nucleic acid comprising a sequence of at least
CC   18 bases of a segment of chemically pretreated DNA of genes associated
CC   with cell signalling. The activity of the modified sequences of the
CC   invention may be described as cytostatic. The object of the invention is
CC   to provide the chemically modified DNA of genes associated with cell
CC   signalling, as well as oligonucleotides and/or PNA-oligomers for
CC   detecting cytosine methylations, as well as a method which is
CC   particularly suitable for the diagnosis and/or therapy of genetic and
CC   epigenetic parameters of genes associated with cell signalling. The
CC   chemically modified DNA provided by the invention is useful for diagnosis
CC   and therapy of diseases such as solid tumours and cancer. The sequences
CC   given in records ABL70111-ABL70626 represent chemically pre-treated
CC   genomic DNA's of genes associated with cell signalling.
CC   Note: The sequence data for this patent is not represented in the printed
CC   specification, but is based on sequence information supplied by the
CC   European Patent Office.
XX
SQ   Sequence 8210 BP; 2370 A; 76 C; 1595 G; 4169 T; 0 other;

  Query Match             2.3%;  Score 52.8;  DB 24;  Length 8210;
  Best Local Similarity   46.3%;  Pred. No. 0.083;
  Matches  210;  Conservative    0;  Mismatches  242;  Indels    2;  Gaps    1;

Qy       1784 ATAAAAAAGAAGTTATACCTGTAACTCATAATTTAACATTGAGAAAAACGGTGACTGGTT 1843
               | |  ||| |  || |  |    |   || ||||||| ||| | ||  |   |  ||  |
Db       4672 AAATTAAACATATTTTTACGTCACTCCACAATTTAAAATTTAAAACCAATCTTTCTCATA 4613

Qy       1844 TAGCTGGTGACAGAACTAAAGATTTCCATTTTGAAATTGAATTAAAAAATAATAAGCAAG 1903
               ||  |  |   | || |||| |  |  ||||| | |   |||| | |||   |||   |
Db       4612 TACTTATTATAAAAAATAAAAACCTTAATTTTTATAAAAAATTTATAAACTTTAATTCAA 4553

Qy       1904 AATTGCTTTCTCAAACTGTTAAAACAGATAAAACAAACCTCGAATTTAAAGATGGTAAAG 1963
```

```
Db       4552 AAATAATTAAATCTCATTTAATACCAAATAAAAAAAACAACAAATAAAAATAAATTAACT 4493

Qy       1964 CAACCATTAATTTAAAACATGGGGAAAGTTTAACACTTCAAGG--TTTACCAGAAGGTTA 2021
Db       4492 AACATATATCATCCTTACCTAAAAAACAAAAAATACCACAATACCTTCACTATAACAAAT 4433

Qy       2022 TTCTTACCTTGTCAAAGAAACAGATTCTGAAGGCTATAAGGTTAAAGTTAATAGCCAAGA 2081
Db       4432 CCTCTAACCTACCAAAACAACCTCCTAATATAACACACATACTAAAACTATAATAAATAA 4373

Qy       2082 AGTAGCAAATGCTACAGTTTCAAAAACAGGAATAACAAGTGATGAGACACTTGCTTTTGA 2141
Db       4372 AAAAATAAAAAACAAAACAAAAAAAAAAAAAAAAAAAAAATAATTAAACATATTAAAATACA 4313

Qy       2142 AAATAATAAAGAGCCTGTTGTTCCTACAGGAGTTGATCAAAAGATCAATGGCTATCTAGC 2201
Db       4312 AAACAAAACAATACTATTCTATCCTACCTAAAACATTAAAAATACGATTCAACTTCTAAC 4253

Qy       2202 TTTGATAGTTATCGCTGGTATCAGTTTGGGGATC 2235
Db       4252 TACCATTCCTTTACTACTTATAAAATTAAAAATC 4219
```

RESULT 12
AAS61282/c
ID AAS61282 standard; DNA; 8210 BP.
XX
AC AAS61282;
XX
DT 29-JAN-2002 (first entry)
XX
DE Human gene regulation-associated gene oligonucleotide #237.
XX
KW Human; Gene regulation-associated gene; severe combined immunodeficiency;
KW cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
KW renal disease; Preeclampsia; cardiac allograft vascular disease;
KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
KW immunostimulant; cardiant; antiinflammatory; coagulant; antiasthmatic;
KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.
XX
OS Homo sapiens.
XX
PN WO200177375-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-EP03968.
XX
PR 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-017470/02.
XX
PT New nucleic acid sequences from chemically modified genes associated
PT with gene regulation, useful for analysing cytosine methylations for
PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
PT disease -
XX
PS Disclosure; SEQ ID No 243; 26pp; English.
XX
CC The invention relates to 224 nucleic acid sequences comprising at least
CC 18 bases of a chemically pretreated gene associated with gene regulation
CC selected from 43 known genes (or complementary sequences). The
CC chemical pretreatment converts cytosine bases unmethylated at the
CC 5-position to uracil or another base with hybridisation behaviour
CC dissimilar to cytosine, to enable analysis of cytosine methylations.
CC The DNA sequences, oligomers (or sets/arrays) and method are
CC useful in the diagnosis of diseases (or predisposition to diseases)
CC associated with gene regulation and in therapy of such diseases, by
CC enabling analysis of the cytosine methylation patterns of such genes,
CC kits are provided. They are especially useful in diagnosis
CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
CC preeclampsia, graft versus-host disease. The present sequence is a
CC sequence included in the sequence data for this specification and is
CC associated with the human gene regulation-associated genes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 8210 BP; 2370 A; 76 C; 1595 G; 4169 T; 0 other;

Query Match 2.3%; Score 52.8; DB 24; Length 8210;
Best Local Similarity 46.3%; Pred. No. 0.083;
Matches 210; Conservative 0; Mismatches 242; Indels 2; Gaps 1;

```
Qy       1784 ATAAAAAAGAAGTTATACCTGTAACTCATAATTTAACATTGAGAAAAACGGTGACTGGTT 1843
Db       4672 AAATTAAACATATTTTTACGTCACTCCACAATTTAAAATTTAAAACCAATCTTTCTCATA 4613

Qy       1844 TAGCTGGTGACAGAACTAAAGATTTCCATTTTGAAATTGAATTAAAAAATAATAAGCAAG 1903
Db       4612 TACTTATTATAAAAAATAAAAACCTTAATTTTTATAAAAAATTTATAAACTTTAATTCAA 4553

Qy       1904 AATTGCTTTCTCAAACTGTTAAAACAGATAAAACAAACCTCGAATTTAAAGATGGTAAAG 1963
Db       4552 AAATAATTAAATCTCATTTAATACCAAATAAAAAAAACAACAAATAAAAATAAATTAACT 4493

Qy       1964 CAACCATTAATTTAAAACATGGGGAAAGTTTAACACTTCAAGG--TTTACCAGAAGGTTA 2021
Db       4492 AACATATATCATCCTTACCTAAAAAACAAAAAATACCACAATACCTTCACTATAACAAAT 4433

Qy       2022 TTCTTACCTTGTCAAAGAAACAGATTCTGAAGGCTATAAGGTTAAAGTTAATAGCCAAGA 2081
Db       4432 CCTCTAACCTACCAAAACAACCTCCTAATATAACACACATACTAAAACTATAATAAATAA 4373

Qy       2082 AGTAGCAAATGCTACAGTTTCAAAAACAGGAATAACAAGTGATGAGACACTTGCTTTTGA 2141
Db       4372 AAAAATAAAAAACAAAACAAAAAAAAAAAAAAAAAAAAAATAATTAAACATATTAAAATACA 4313

Qy       2142 AAATAATAAAGAGCCTGTTGTTCCTACAGGAGTTGATCAAAAGATCAATGGCTATCTAGC 2201
Db       4312 AAACAAAACAATACTATTCTATCCTACCTAAAACATTAAAAATACGATTCAACTTCTAAC 4253

Qy       2202 TTTGATAGTTATCGCTGGTATCAGTTTGGGGATC 2235
Db       4252 TACCATTCCTTTACTACTTATAAAATTAAAAATC 4219
```

RESULT 13
ABK31380/c
ID ABK31380 standard; DNA; 8210 BP.
XX
AC ABK31380;
XX
DT 23-APR-2002 (first entry)
XX
DE Signal transduction associated gene modified DNA #112.
XX
KW Human; signal transduction associated gene; cytosine methylation state;
KW CpG island; signal transduction associated disease; solid tumour; cancer;
KW antitumour; cytostatic; mutant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX

```
PN   WO200200926-A2.
XX
PD   03-JAN-2002.
XX
PF   29-JUN-2001; 2001WO-EP07472.
XX
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-147896/19.
XX
PT   Oligonucleotide for diagnosis and therapy of diseases associated with
PT   signal transduction e.g. cancer, comprises chemically modified genomic
PT   sequences of genes associated with signal transduction  -
XX
PS   Claim 1; SEQ ID No 223; 24pp; English.
XX
CC   The present invention relates to chemically modified DNA sequences of
CC   signal transduction associated genes. The DNA sequences are chemically
CC   modified using a solution of bisulphite, hydrogen sulphite or
CC   disulphite. Also disclosed are oligonucleotides and/or PNA oligomers
CC   for detecting the cytosine methylation state (CpG islands) of these
CC   genes, and a method for the diagnosis and/or therapy of genetic and
CC   epigenetic parameters of genes associated with signal transduction.
CC   The genomic DNA can be obtained from cells or cellular components which
CC   contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,
CC   cerebral-spinal fluid, tissue embedded in paraffin such as tissue from
CC   eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,
CC   histologic object slides, and all their possible combinations. The
CC   sequences of the invention are useful for the diagnosis and therapy of
CC   diseases associated with signal transduction e.g. solid tumours and
CC   cancer. ABK31158-ABK31545 represent chemically pretreated genomic DNA
CC   sequences of different genes associated with signal transduction, or
CC   their complementary sequences.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   European Patent Office.
XX
SQ   Sequence 8210 BP; 2370 A; 76 C; 1595 G; 4169 T; 0 other;

  Query Match             2.3%;  Score 52.8;  DB 24;  Length 8210;
  Best Local Similarity   46.3%;  Pred. No. 0.083;
  Matches  210;  Conservative    0;  Mismatches  242;  Indels    2;  Gaps    1;

Qy       1784 ATAAAAAAGAAGTTATACCTGTAACTCATAATTTAACATTGAGAAAAACGGTGACTGGTT 1843
             | |  ||| |  || |  |   |   || |||||||| ||| | ||  |   |  ||  |
Db       4672 AAATTAAACATATTTTTACGTCACTCCACAATTTAAAATTTAAAACCAATCTTTCTCATA 4613

Qy       1844 TAGCTGGTGACAGAACTAAAGATTTCCATTTTGAAATTGAATTAAAAAATAATAAGCAAG 1903
             ||  |  |   | || |||| |  |  ||||| | |   |||| | |||   |||   |
Db       4612 TACTTATTATAAAAAATAAAAACCTTAATTTTTATAAAAAATTTATAAACTTTAATTCAA 4553

Qy       1904 AATTGCTTTCTCAAACTGTTAAAACAGATAAAACAAACCTCGAATTTAAAGATGGTAAAG 1963
             || |  ||      | | | | || |||||| |||| | |||  ||| |   |||
Db       4552 AAATAATTAAATCTCATTTAATACCAAATAAAAAAAACAACAAATAAAAATAAATTAACT 4493

Qy       1964 CAACCATTAATTTAAAACATGGGGAAAGTTTAACACTTCAAGG--TTTACCAGAAGGTTA 2021
               |   ||    |   || |    ||     || ||  |||    || || | ||
Db       4492 AACATATATCATCCTTACCTAAAAAACAAAAAATACCACAATACCTTCACTATAACAAAT 4433

Qy       2022 TTCTTACCTTGTCAAAGAAACAGATTCTGAAGGCTATAAGGTTAAAGTTAATAGCCAAGA 2081
                || | |  |||  |||    |   |   |    |   ||||  ||  |   |  |
Db       4432 CCTCTAACCTACCAAAACAACCTCCTAATATAACACACATACTAAAACTATAATAAATAA 4373

Qy       2082 AGTAGCAAATGCTACAGTTTCAAAAACAGGAATAACAAGTGATGAGACACTTGCTTTTGA 2141
             |  |  |||    | |    |||||  |  || || ||    | |     ||    |  |
Db       4372 AAAAATAAAAAACAAAACAAAAAAAAAAAAAAAAAAAAATAATTAAACATATTAAAATACA 4313

Qy       2142 AAATAATAAAGAGCCTGTTGTTCCTACAGGAGTTGATCAAAAGATCAATGGCTATCTAGC 2201
```

```
             ||| || | |   |   |  ||||||  |    | ||||  |  | |     |||| |
Db       4312 AAACAAAACAATACTATTCTATCCTACCTAAAACATTAAAAATACGATTCAACTTCTAAC 4253

Qy       2202 TTTGATAGTTATCGCTGGTATCAGTTTGGGGATC 2235
              |   ||   | |      ||| |  ||    |||
Db       4252 TACCATTCCTTTACTACTTATAAAATTAAAAATC 4219


RESULT 14
AAS46429/c
ID   AAS46429 standard; DNA; 6106 BP.
XX
AC   AAS46429;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Tumour suppressor gene derived chemically modified sequence #151.
XX
KW   Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW   cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW   cytosine methylation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200168912-A2.
XX
PD   20-SEP-2001.
XX
PF   15-MAR-2001; 2001WO-EP02955.
XX
PR   15-MAR-2000; 2000DE-1013847.
PR   06-APR-2000; 2000DE-1019058.
PR   07-APR-2000; 2000DE-1019173.
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-602752/68.
XX
PT   Fragments of chemically modified genes associated with tumour suppressor
PT   genes and oncogenes, useful in designing primers and probes for
PT   analysing diseases associated with cytosine methylation state e.g.
PT   cancer  -
XX
PS   Claim 1; SEQ ID No 151; 27pp; English.
XX
CC   The invention relates to a nucleic acid comprising a sequence of 18
CC   bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC   bisulphite, of genes associated with tumour suppression and
CC   oncogenes having a sequence taken from 536 (actually 533 since
CC   numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC   (Ss) and sequences complementary to (Ss). The nucleic acid may be a
CC   peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC   form part of a set of probes for detecting the cytosine methylation state
CC   and/or single nucleotide polymorphisms and also to be used in an
CC   array for analysing diseases associated with CpG dinucleotides e.g.
CC   cancers and tumours. The probes can also be used in a method for
CC   ascertaining genetic and/or epigenetic parameters for the diagnosis
CC   and/or therapy of existing diseases or the predisposition to specific
CC   diseases, by analysing cytosine methylations. The parameters may be
CC   compared to another set of genetic and/or epigenetic parameters, the
CC   differences serving as basis for diagnosis and/or prognosis events which
CC   are disadvantageous to patients. The present sequence is one of the
CC   533 genomic sequences derived from tumour suppressor genes and
CC   oncogenes.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
```

SQ Sequence 6106 BP; 1938 A; 30 C; 849 G; 3289 T; 0 other;

Query Match 2.3%; Score 52.4; DB 22; Length 6106;
Best Local Similarity 46.7%; Pred. No. 0.094;
Matches 271; Conservative 0; Mismatches 301; Indels 8; Gaps 3;

```
Qy       1544 TATATTATTTCACTGATAGTGCTGAATTAGATAAGGATAAACTAAAAGACTATCATGGTT 1603
Db       2939 TATAATAATTATATATAATTAATTTATAATAATTATATATAATTAATTTATAATAATTAT 2880

Qy       1604 TTGGAGACATGAATGATAGTACTTTAGCAGTTGCTAAAATCCTTGTAGAATACGCTCAAG 1663
Db       2879 ATAATCAATATAATATAATATATTTAATATTTAATATAATATATTTAATATATAATTATA 2820

Qy       1664 ATAGTAATCCTCCACAGCTAACTGACCTTGATTTCTTTATTCCGAATAACAATAAATATC 1723
Db       2819 TTAAT-ATAATATATAATTATATTATATTAATTTATATTAATATATAAATAATACATATA 2761

Qy       1724 AATCTCTTATTGGAACTCAGTGGCATCCAGAAGATTTAGTTGATATTATTCGTATGGAAG 1783
Db       2760 ATTATATTATATATAATTAATTATATATATAACATTTAATATATAATATATTAATATAAT 2701

Qy       1784 ATAAAAAAGAAGTTATACCTGTAACTCATAATTTAACATTGAGAAAAACGGTGACTGGTT 1843
Db       2700 ATATTAAAAAACTAAAATATAAACATTAATAACTATTTT-----AATAATTTCTCTTTAA 2646

Qy       1844 TAGCTGGTGACAGAACTAAAGATTTCCATTTTGAAATTGAATTAAAAAATAATAAGCAAG 1903
Db       2645 CTCTTAAAAACAACAAAAACCATTAAAATAAAAACATTTCATCATAAAAAAAACAATATT 2586

Qy       1904 AATTGCTTTCTCAAACTGTTAAAACAGATAAAACAAACCTCGAATTTAAAGATGGTAAAG 1963
Db       2585 AACTAATTAATTAATTACTTTATCACACTAATTATCTTATTAAAAATAAACTATAAAAAA 2526

Qy       1964 CAACCATTAATTTAAAACATGGGGAAAGTTTAACACTTCAAGGTTTACCAGAAGGTTATT 2023
Db       2525 CAAAAACAAAAATCAAAAAAACCAATCAAAAAACTATTACAACAATTCCAATAAAAAATA 2466

Qy       2024 CTTACCTTGTCAAAGAAACAGATTCTGAAGGCTATAAGGTTAAAGTTAATAGCCAAGAAG 2083
Db       2465 ATAACTTAATCAAAAAAACTTATTAACTAACCAATTA--CTAAAAAAAAACTAAAAAAAA 2408

Qy       2084 TAGCAAATGCTACAGTTTCAAAAACAGGAATAACAAGTGA 2123
Db       2407 TAAAAAATAAAAAAATTTTAAAACGATCTATATAAAATAA 2368
```

RESULT 15
ABK40031/c
ID ABK40031 standard; DNA; 6106 BP.
XX
AC ABK40031;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human chemically pretreated gene sequence #57 strand 1.
XX
KW Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;
KW cytostatic; ALDH6; CYP11A; CYP11B1; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;
KW UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
PN WO200202806-A2.
XX
PD 10-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP07470.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-154757/20.
XX
PT New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
PT useful for detecting cytosine methylation state of genes associated
PT with pharmacogenomics and for therapy of diseases e.g. cancer -
XX
PS Claim 1; SEQ ID No 113; 24pp; English.
XX
CC The invention relates to a nucleic acid comprising a sequence at
CC least 18 bases in length of a segment of the chemically pretreated DNA
CC of genes associated with pharmacogenomics according to one of the
CC sequences of the genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B1
CC (NM_000497), CYP3A3 (NM_000776 and NM_017460), DPYD (NM_000110), EPHX2
CC (NM_001979), OCLN (NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360),
CC MRP (NM_004996, NM_019900, NM_019901, NM_019902, NM_019862, NM_019898,
CC NM_019899) and their complementary sequences, or a sequence (S1) chosen
CC from 87 sequences and their complements. The chemical pretreatment
CC is bisulphite treatment to convert cytosines (but not methyl-cytosines)
CC into uracils. Also included are an oligomer (II) in particular an
CC oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in
CC each case at least one base sequence having a length of 9 nucleotides
CC which hybridises to or is identical to a chemically pretreated DNA of
CC genes associated with pharmacogenomics and their complements, arranged in
CC an array for analysing diseases associated with the methylation state
CC (CpG) and/or detecting SNPs (single nucleotide polymorphisms)
CC of the 87 sequences. The oligomers may also be used as PCR primers.
CC The set of 87 nucleic acids and their complements is useful for diagnosis
CC and therapy of solid tumours and cancer. The present sequence
CC represents one the 87 DNA sequences or its complement.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 6106 BP; 1938 A; 30 C; 849 G; 3289 T; 0 other;

Query Match 2.3%; Score 52.4; DB 24; Length 6106;
Best Local Similarity 46.7%; Pred. No. 0.094;
Matches 271; Conservative 0; Mismatches 301; Indels 8; Gaps 3;

```
Qy       1544 TATATTATTTCACTGATAGTGCTGAATTAGATAAGGATAAACTAAAAGACTATCATGGTT 1603
Db       2939 TATAATAATTATATATAATTAATTTATAATAATTATATATAATTAATTTATAATAATTAT 2880

Qy       1604 TTGGAGACATGAATGATAGTACTTTAGCAGTTGCTAAAATCCTTGTAGAATACGCTCAAG 1663
Db       2879 ATAATCAATATAATATAATATATTTAATATTTAATATAATATATTTAATATATAATTATA 2820

Qy       1664 ATAGTAATCCTCCACAGCTAACTGACCTTGATTTCTTTATTCCGAATAACAATAAATATC 1723
Db       2819 TTAAT-ATAATATATAATTATATTATATTAATTTATATTAATATATAAATAATACATATA 2761

Qy       1724 AATCTCTTATTGGAACTCAGTGGCATCCAGAAGATTTAGTTGATATTATTCGTATGGAAG 1783
Db       2760 ATTATATTATATATAATTAATTATATATATAACATTTAATATATAATATATTAATATAAT 2701

Qy       1784 ATAAAAAAGAAGTTATACCTGTAACTCATAATTTAACATTGAGAAAAACGGTGACTGGTT 1843
Db       2700 ATATTAAAAAACTAAAATATAAACATTAATAACTATTTT-----AATAATTTCTCTTTAA 2646

Qy       1844 TAGCTGGTGACAGAACTAAAGATTTCCATTTTGAAATTGAATTAAAAAATAATAAGCAAG 1903
Db       2645 CTCTTAAAAACAACAAAAACCATTAAAATAAAAACATTTCATCATAAAAAAAACAATATT 2586

Qy       1904 AATTGCTTTCTCAAACTGTTAAAACAGATAAAACAAACCTCGAATTTAAAGATGGTAAAG 1963
Db       2585 AACTAATTAATTAATTACTTTATCACACTAATTATCTTATTAAAAATAAACTATAAAAAA 2526

Qy       1964 CAACCATTAATTTAAAACATGGGGAAAGTTTAACACTTCAAGGTTTACCAGAAGGTTATT 2023
```

```
Db       2525 CAAAAACAAAAATCAAAAAAACCAATCAAAAAACTATTACAACAATTCCAATAAAAAATA 2466

Qy       2024 CTTACCTTGTCAAAGAAACAGATTCTGAAGGCTATAAGGTTAAAGTTAATAGCCAAGAAG 2083
              | || |  ||||| |||| |||    |  | || |   ||||   ||     || ||
Db       2465 ATAACTTAATCAAAAAAACTTATTAACTAACCAATTA--CTAAAAAAAAACTAAAAAAAA 2408

Qy       2084 TAGCAAATGCTACAGTTTCAAAAACAGGAATAACAAGTGA 2123
              ||  ||||   | | ||| ||||  |   |||  || | |
Db       2407 TAAAAAATAAAAAAATTTTAAAACGATCTATATAAAATAA 2368
```

Search completed: August 19, 2003, 09:47:38
Job time : 597 secs